Case Report

Predicting product quality and optimising process design using dynamic time warping in batch processes with varying batch times

Shuto Yamakage [a], Kazutoshi Terauchi [b], Fumiya Hamada [b], Toshinori Yamaji [b], Hiromasa Kaneko [a,*]

[a] Department of Applied Chemistry, School of Science and Technology, Meiji University, 1-1-1 Higashi-Mita, Tama-ku, Kawasaki, Kanagawa, 214-8571, Japan
[b] Core Process Engineering Development Group, Production Engineering Department, DIC Corporation, 631 Sakado, Sakura, Chiba, 285-8668, Japan

ARTICLE INFO

ABSTRACT

In this study, the proposed method combines dynamic time warping with a genetic algorithm, enabling the construction of a robust machine learning model tailored for batch process data characterised by diverse batch and sampling times. This comprehensive model incorporates crucial factors, including initial process conditions, time-series data of process variables, and product quality. Furthermore, it allows for the simultaneous design of both initial process conditions and time-series data of the process variables (that is, batch profiles) for achieving the target values of predefined product quality. By leveraging actual semi-batch process data for polymer production, we confirmed that the proposed method can accurately predict the product quality during and at the end of the batch process. Concurrently, it facilitates the design of reasonable initial process conditions and batch profiles.

1. Introduction

The utilisation of machine learning (ML) models or soft sensors, represented by $y = f(x)$, plays a crucial role in various industrial processes and equipment, particularly in scenarios where certain process variables y are difficult to measure but can be estimated through the measurement of easily accessible process variables x [1]. The constructed soft sensor enables the prediction of y-values from x-values, with the predicted y-values used instead of the measured y-values for enhanced process control. In a batch process, the product quality at the end of each batch or endpoint can be predicted before the batch is operated employing y as the product quality or properties, and x as the batch time and initial process conditions [2]. By inverse analysis of the soft sensor model, the batch time and initial process conditions can be designed from the target values of y. Furthermore, by adding time-series data of process variables, that is, batch profiles, to x, the product quality or properties can be predicted during batch operation, and the batch profiles can be designed from the target values of y [3].

In batch process design and optimisation, various studies have explored batch-to-batch optimisation for operating and control conditions [4–7], within-batch optimisation for batch processes by predicting the product quality and properties of materials during batch operation [8–10], and batch process control to adhere to the batch profile of a process variable [11–15]. In addition, end-point prediction [16,17] and design of experiments [18–20] make it possible to optimise the process conditions of the batch process. When x contains data for each time for each process variable, if the batch time is the same for each batch, there is no problem because the process variables in x are aligned for all samples or batches. However, in cases where batch data exhibit different batch times, the model cannot be simply constructed since the values of the process variables are not aligned. To construct a model for such a dataset, time-series data with short batch times were complemented, and a model was constructed after Fourier transforming the time-series data [3].

This study focuses on dynamic time warping (DTW) [21], a method that can calculate the similarity between samples of time-series data at different times and attempts to predict endpoints y and design batch processes using DTW. The datasets handled in this study are time-series data containing continuous values of process variables in a batch process, and the length of each time-series data differs owing to different

* Corresponding author.
E-mail address: hkaneko@meiji.ac.jp (H. Kaneko).

sampling and batch times. Therefore, DTW was used instead of the other distances. In the training data, we calculated the DTW between all samples and added the DTW values to x, in addition to batch times and initial process conditions. In the test data, the DTW with all batches of training data for time-series data was calculated and used as x to predict the y-values. Furthermore, to design and search for the time-series data of process variables, that is, batch profiles in which y has target values, a genetic algorithm (GA) [22] was used, and the generation of virtual batch profiles and the prediction of y were repeated. The effectiveness of the proposed method was verified using data measured during an actual semi-batch process for polymer production.

2. Method

2.1. Dynamic time warping (DTW)

Because the dataset used in this study involves time-series data containing continuous values of process variables in a batch process, and the length of each time-series data differs owing to different sampling times and batch times, DTW was introduced to calculate the similarity between samples in a batch process. DTW calculates the similarity between two sets of time-series data. We denote the time series data T by t_1, $t_2, \ldots, t_p$ and other time series data U by $u_1, u_2, \ldots, u_q$. Let $\delta(i,j) = |t_i - u_j|$ be the distance between each element i of T and j of U, and then, the similarity $S_{\mathrm{T,U}}$ between T and U is calculated as follows:

$$S(\mathrm{T}, \mathrm{U}) = \min_{\mathrm{w}} \sum_{k=1}^{r} \delta(w_k) \qquad (1)$$

where w is the minimum distance combination of i and j and r is the number of combinations. In this study, dtwalign [23] was used to calculate the DTW.

2.2. Endpoint prediction and batch profile design with DTW and GA

The features of the training data were divided into the initial process conditions of the batch process, batch times, time-series data of process variables, and endpoint properties. The endpoint properties were set as y and the initial process conditions of the batch process were set as x. For the time-series data, the DTW was calculated for each process variable between all samples in the training data. The DTW values for each sample of training data and each process variable were added to x. The number of initial process conditions was k, the number of process variables was m, the number of samples in the training data was n, and the number of x variables was $k + mn$. An ML model $y = f(x)$ was constructed between x and y using a regression analysis.

For the test data or new data, the initial process conditions and DTW calculated between the training data and test data in the time series data were input into the model as x to predict the y-values. Furthermore, the initial process conditions and time-series data of the process variables, that is batch profiles, were proposed to obtain the target y-values. A GA was used to generate data including values of initial conditions and time-series data simultaneously with which the predicted y-values meet the target y-values. The initial process conditions and batch profiles are represented as chromosomes in the GA, and the inverse of the absolute value of the difference between the predicted and target y-values is the fitness of the GA. The chromosomes of the initial process conditions and batch profiles can then be output such that the fitness increases, that is, the predicted y-values are close to the target y-values.

3. Results and discussion

The effectiveness of the proposed method was validated using a dataset obtained in an actual semi-batch process of polymer production. In the semi-batch process, multiple types of monomers and initiators are used to polymerise materials. Two variables, y_1 and y_2, indicate the

Table 1
Prediction results of DCV for y_1.

	AV	FV	ST	DTW
r^2	0.637	0.504	0.526	0.494
RMSE	0.92	1.03	1.25	0.96

Table 2
Prediction results of DCV for y_2.

	AV	FV	ST	DTW
r^2	0.195	0.229	0.438	0.762
RMSE	428	390	350	223

properties of the polymer material. For the x variables, there are 11 initial process conditions, and there are three process variables: reaction solution temperature T_r, jacket temperature T_j, and drop volume F. The semi-batch process was divided into two parts. The first part is the dropping of monomers and the polymerisation initiator, followed by completion of the reaction. In the second part, the liquid temperature was maintained, and y_1 and y_2 were measured four to eight times. Here, F exists only in the first part. The number of batches was 11, and the number of samples was 60 because each batch was sampled several times.

Using the average of each process variable (AV) and the final value of each process variable (FV) and by adjusting the time-series data to the shortest batch time (ST), each sample of time-series data was prepared for comparison with the proposed DTW-based method. Partial least squares regression [24], ridge regression [25], least absolute shrinkage and selection operator [26], elastic net [27], support vector regression [28], random forests [29], and Gaussian process regression [30] were used as the regression analysis methods, and the method with the best predictive ability was selected for each time-series data processing method. Because the number of batches was small in this study, double cross-validation (DCV) [31] was used to evaluate the predictive accuracy.

The prediction results for the DCV are shown in Tables 1 and 2 and Figs. 1 and 2 r^2 is the determinant coefficient, and RMSE is the root-mean-square error of the prediction. From Table 1, the r^2 of AV was the largest for y_1; however, from Fig. 1, there were no significant changes in the prediction results for the four methods. y_1 does not depend much on the batch time and batch profile of the process variables and is determined by the initial process conditions. However, the r^2 of DTW was the largest for y_2, and Fig. 2 shows that only DTW could predict large y_2-values accurately. y_2 depends not only on the initial process conditions, but also on the batch profiles. It was confirmed that the proposed DTW could accurately predict the physical properties during and at the end of the batch.

A GA-based inverse analysis was performed using the constructed y_1 and y_2 prediction models. The target values of y_1 and $\underline{y}_2$ were set as 50 and 1523, respectively. The time for the first part was 305 min, and the candidates for the second part ranged from 30 to 380 min. The GA optimises the time of the second part, initial process conditions, and batch profiles of T_r, T_j, and F to reduce the distance between the predicted and target y-values. T_r and T_j were assumed to have the same value from 100 to 140 °C, and F was set as the minimum to maximum value of the training data separated by 5 min as the upper and lower limits, respectively. The number of chromosomes was set to 100, the number of generations to 50, the crossover rate to 0.5, and the mutation rate to 0.3. The weight of y_1 was set to 2^{-2}, 2^{-1}, $\ldots$, 2^7, 2^8, and the weight of y_2 was set to 1 for the distance between the predicted and target y-values, and then, GA was conducted.

The results of the inverse analysis based on the GA are shown in Fig. 3. The initial process conditions and batch profiles were predicted to be within the specified ranges for both y_1 and y_2 by changing the weight
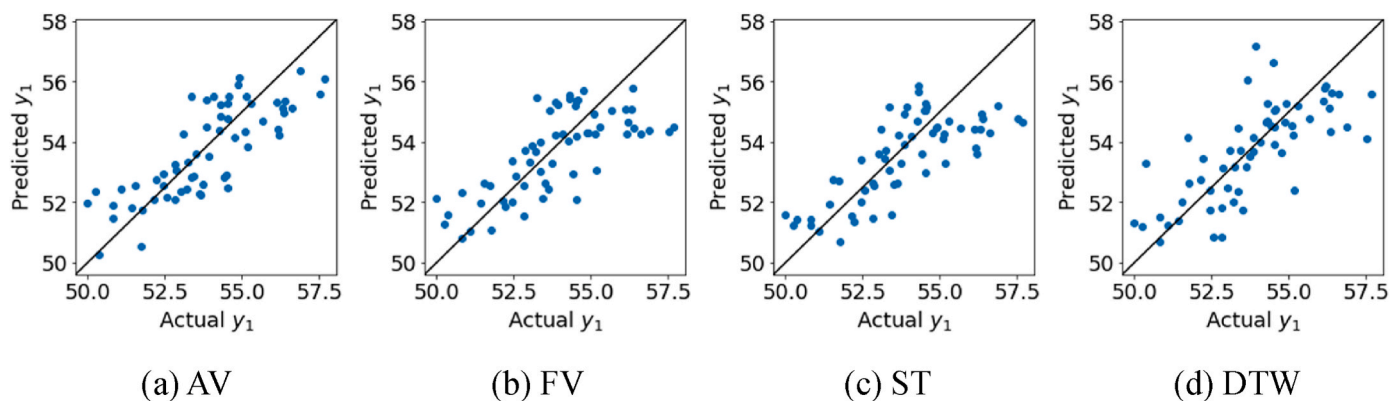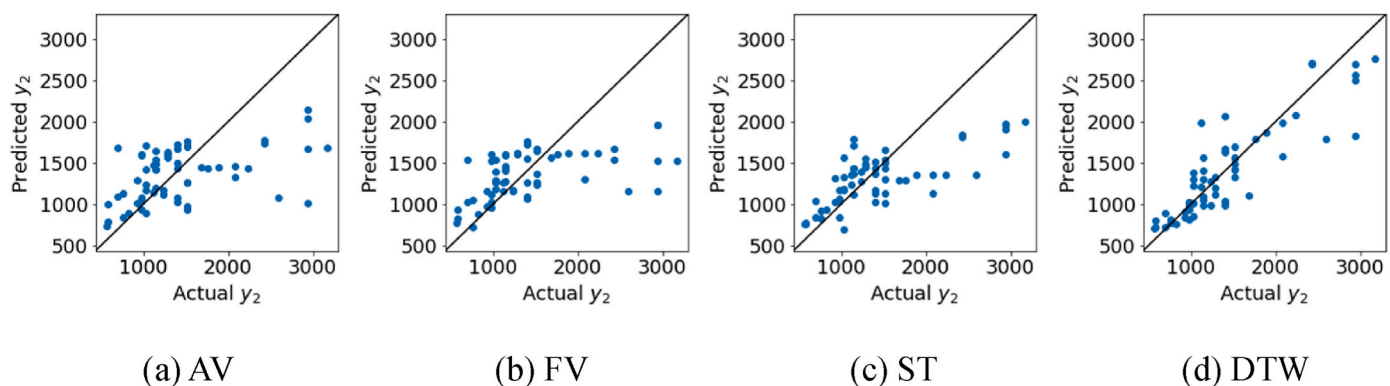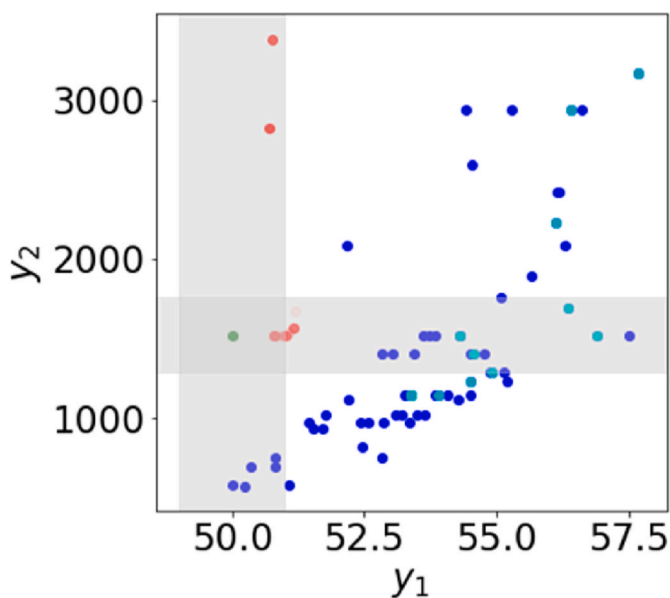
(a) AV (b) FV (c) ST (d) DTW

Fig. 1. Actual y_1 vs predicted y_1 in DCV.

(a) AV (b) FV (c) ST (d) DTW

Fig. 2. Actual y_2 vs predicted y_2 in DCV.

Fig. 3. Results of inverse analysis of the models. The dark blue, light blue, green, and red points indicate samples measured in the middle of the batch in the training data, samples measured at endpoint in the training data, the target value, and the predicted values proposed by GA. The grey area indicates the target ranges.

of y_1. Fig. 4 shows the time-series data for the process variables proposed in the GA-based inverse analysis. For the time-series data of the process variables that were considered to meet the product specifications, the

time of the second part was 30 min, which was much shorter than the time required for the training data. *F*, which changed over time, had a reasonable time variation compared with the variations in the training data. Operating with these time variations of the process variables improves the polymer properties y_1 and y_2 and achieves the target ranges. The proposed method could appropriately design the batch time and batch profiles during the process, in addition to the initial process conditions of the batch process.

4. Conclusions

For batch processes with different batch times, we proposed a method for constructing property prediction models using DTW and ML, and a method for designing initial process conditions, batch time, and batch profiles to achieve target values of product quality using GA. The similarity between samples of time-series data for each process variable was calculated using DTW and was used as feature x of the batch process. After the models to predict properties y during the process were constructed, the initial process conditions, batch time, and batch profiles were designed by inverse analysis of the models using the GA. Verification using an actual semi-batch process dataset confirmed that the proposed method can accurately predict the two properties of polymer materials in the batch process and can design the initial process conditions, batch time, and batch profiles that satisfy the specifications of both properties, which were not available in the existing dataset.

The proposed method makes it possible to construct an ML model to predict product quality in a batch process and optimise the batch process by inverse analysis of the model, which contributes to improving the operating efficiency and reducing the number of out-of-specification products. Using the proposed method, the time evolution of the process variables to achieve the target product quality can be predicted at
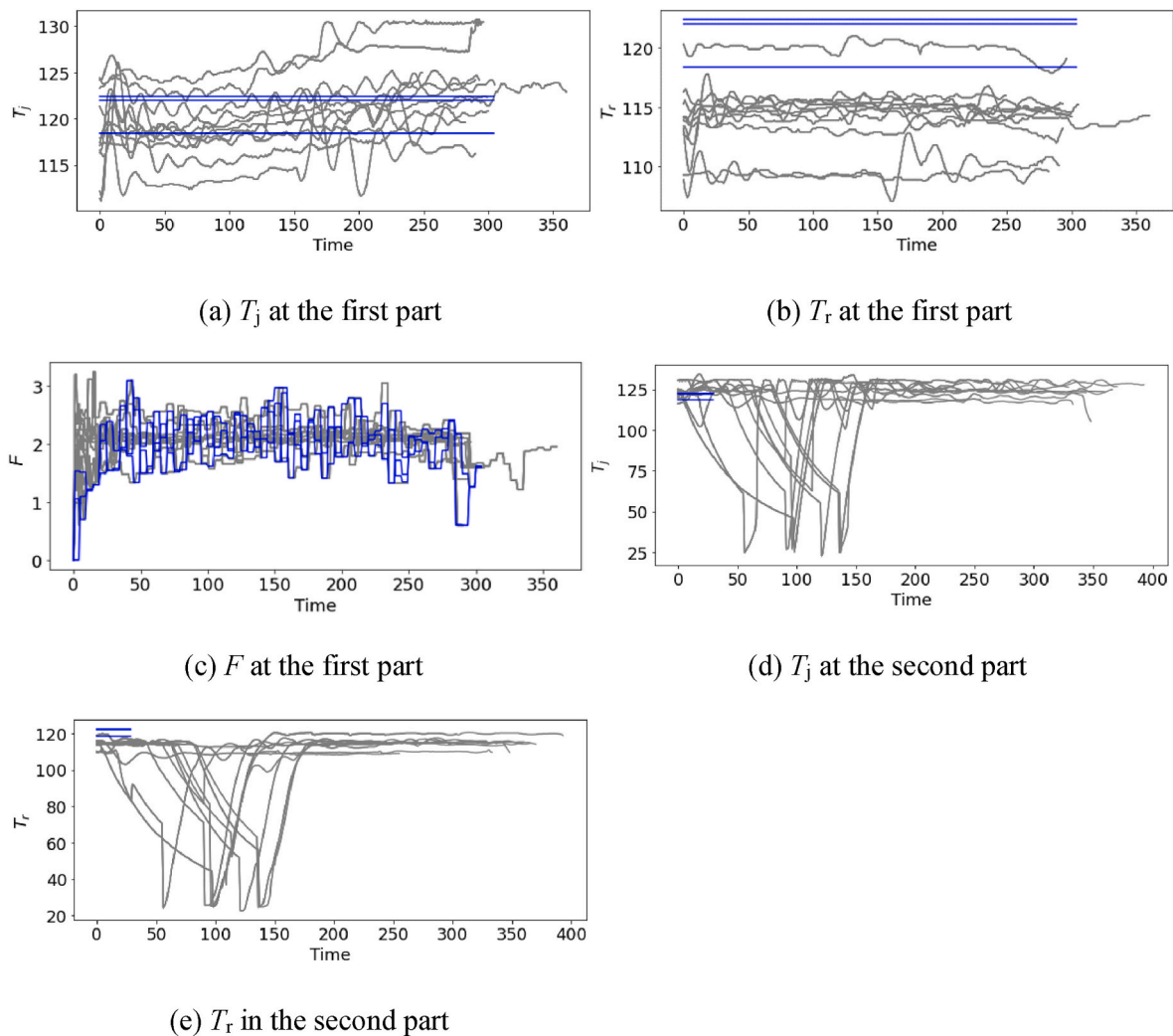
(a) T_j at the first part

(b) T_r at the first part

(c) F at the first part

(d) T_j at the second part

(e) T_r in the second part

Fig. 4. Batch profiles designed with the proposed method. Blue and grey lines indicate designed batch profiles and batch profiles in the training data.

the start of the batch process operation. Furthermore, because the proposed method can estimate the values of product quality during online batch processing, it is possible to optimise the process conditions and variables of the batch process in real time, which also enables feedback control to achieve the target values of product quality. Future research should include process control based on real-time prediction of the forward time evolution of process variables in batch processes.

CRediT authorship contribution statement

Shuto Yamakage: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Kazutoshi Terauchi:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Fumiya Hamada:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Toshinori Yamaji:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Hiromasa Kaneko: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The data that has been used is confidential.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cscee.2024.100655.

References

[1] Y. Jiang, S. Yin, J. Dong, O. Kaynak, A review on soft sensors for monitoring, control, and optimization of industrial processes, IEEE Sensor. J. 21 (11) (2021) 12868–12881. https://10.1109/JSEN.2020.3033153.

[2] V. Brunner, M. Siegl, D. Geier, T. Becker, Challenges in the development of soft sensors for bioprocesses: a critical review, Front. Bioeng. Biotechnol. 9 (2021) 722202. https://10.3389/fbioe.2021.722202.

[3] H. Kaneko, Direct prediction of the batch time and process variable profiles using batch process data based on different batch times, Comput. Chem. Eng. 169 (2022) 108072.

[4] R. Jia, Z. Mao, F. Wang, D. He, Batch-to-batch optimization of cobalt oxalate synthesis process using modifier-adaptation strategy with latent variable model, Chemometr. Intell. Lab. Syst. 140 (2015) 73–85.

[5] R. Hille, H.M. Budman, Batch-to-Batch optimization of chemical processes in the presence of model-plant mismatch using a variable subset of model parameters, IFAC-PapersOnLine 49 (2016) 61–66.

[6] R. Jia, Z. Mao, D. He, F. Chu, Hierarchical batch-to-batch optimization of cobalt oxalate synthesis process based on data-driven model, Chem. Eng. Res. Des. 144 (2019) 185–197.

[7] F. Chu, J. Wang, Y. Wang, G. Chai, R. Jia, F. Wang, Can. Integrated operation optimization strategy for batch process based on process transfer model under disturbance, J. Chem. Eng. (2022) 1, https://doi.org/10.1002/cjce.24362.

[8] J. Zhang, A reliable neural network model based optimal control strategy for a batch polymerization reactor, Ind. Eng. Chem. Res. 43 (4) (2004) 1030–1038.

[9] Y. Liu, Z. Zhang, J. Chen, Ensemble local kernel learning for online prediction of distributed product outputs in chemical processes, Chem. Eng. Sci. 137 (2015) 140–151.

[10] E.S. Yadav, P. Shettigar J, S. Poojary, S. Chokkadi, G. Jeppu, T. Indiran, Data-Driven modeling of a pilot plant batch reactor and validation of a nonlinear model predictive controller for dynamic temperature profile tracking, ACS Omega 6 (26) (2021) 16714–16721.

[11] J. Camacho, D. Lauri, B. Lennox, M. Escabias, M. Valderrama, Evaluation of smoothing techniques in the run to run optimization of fed-batch processes with u-PLS, J. Chemom. 29 (2015) 338–348. https://10.1002/cem.2711.

[12] P. Shettigar J, J. Kumbhare, E.S. Yadav, T. Indiran, Wiener-Neural-Network-based modeling and validation of generalized predictive control on a laboratory-scale batch reactor, ACS Omega 7 (2022) 16341–16351.

[13] R. Kamesh, K.Y. Rani, Nonlinear control strategies based on Adaptive ANN models: multi-product semi-batch polymerization reactor case study, Chem. Eng. Res. Des. 121 (2017) 255–274.

[14] R. Nian, J. Liu, B. Huang, A review on reinforcement learning: introduction and applications in industrial process control, Comput. Chem. Eng. 139 (2020) 106886.

[15] S. Teworte, K. Malcı, L.E. Walls, M. Halim, L. Rios-Solis, Recent advances in fed-batch microscale bioreactor design, Biotechnol. Adv. 55 (2022) 107888.

[16] J. Mareš, J. Kukal, P. Hrnčiřík, J. Náhlík, Process state classification of fed-batch fermentation based on process variables analysis, Biochem. Eng. J. 112 (2016) 178–185.

[17] G. Gins, P.V. den Kerkhof, J.F.M.V. Impe, Hybrid derivative dynamic time warping for online industrial batch-end quality estimation, Ind. Eng. Chem. Res. 51 (2012) 6071–6084.

[18] B. Szilagyi, A. Eren, J.L. Quon, C.D. Papageorgiou, Z.K. Nagy, Application of model-free and model-based quality-by-control (QbC) for the efficient design of pharmaceutical crystallization processes, Cryst. Growth Des. 20 (2020) 3979–3996.

[19] A.S. Okouzi, A.O.A. Ibhadode, A.I. Obanor, Response surface methodology (RSM) optimization of the batch process in a rectangular passive greenhouse dryer, JERA 56 (2021) 145–161.

[20] M. Shanmugaprakash, S. Venkatachalam, K. Rajendran, A. Pugazhendhi, Biosorptive removal of Zn(II) ions by Pongamia oil cake (Pongamia pinnata) in batch and fixed-bed column studies using response surface methodology and artificial neural network, J. Environ. Manag. 227 (2018) 216–228.

[21] K. Qiu, J. Wang, R. Wang, Y. Guo, L. Zhao, Soft sensor development based on kernel dynamic time warping and a relevant vector machine for unequal-length batch processes, Expert Ayat. Appl. 182 (2021) 115223.

[22] S. Liu, W. Sun, Attention mechanism-aided data- and knowledge-driven soft sensors for predicting blast furnace gas generation, Energy 262 (2023) 125498.

[23] https://github.com/statefb/dtwalign (accessed 11 December 2022).

[24] W. Shao, W. Han, Y. Li, Z. Ge, D. Zhao, Enhancing the reliability and accuracy of data-driven dynamic soft sensor based on selective dynamic partial least squares models, Control Eng. Pract. 127 (2022) 105292.

[25] Z. He, J. Qian, J. Li, M. Hong, Y. Man, Data-driven soft sensors of papermaking process and its application to cleaner production with multi-objective optimization, J. Clean. Prod. 372 (2022) 133803.

[26] Y. Liu, X. Yu, J. Zhao, C. Pan, K. Sun, Development of a robust data-driven soft sensor for multivariate industrial processes with non-Gaussian noise and outliers, Mathematics 10 (2022) 3837, https://doi.org/10.3390/math10203837.

[27] F. Zhang, K. Sun, X. Wu, A novel variable selection algorithm for multi-layer perceptron with elastic net, Neurocomputing 361 (2019) 110–118.

[28] Y. Dai, A. Liu, M. Chen, Y. Liu, Y. Yao, Enhanced soft sensor with qualified augmented samples for quality prediction of the polyethylene process, Polymers 14 (2022) 4769, https://doi.org/10.3390/polym14214769.

[29] B. Szelag, J. Drewnowski, G. Łagód, D. Majerek, E. Dacewicz, F. Fatone, Soft sensor application in identification of the activated sludge bulking considering the technological and economical aspects of smart systems functioning, Sensors 20 (2020) 1941, https://doi.org/10.3390/s20071941.

[30] Y. Zhang, H. Jin, H. Liu, B. Yang, S. Dong, Deep semi-supervised ust-in-time learning based soft sensor for mooney viscosity estimation in industrial rubber mixing process, Polymers 14 (2022) 1018, https://doi.org/10.3390/polym14051018.

[31] T. Dias, R. Oliveira, P.M. Saraiva, M.S. Reis, Linear and non-linear soft sensors for predicting the research octane number (RON) through integrated synchronization, resolution selection and modelling, Sensors 22 (2022) 3734, https://doi.org/10.3390/s22103734.